# *Comparison of Chinese Gaokao and Western university undergraduate admission criteria: Australian ATAR as an example*

# Comparison of Chinese Gaokao and Western university undergraduate admission criteria: Australian ATAR as an example

Alan Farley

Victoria University, Melbourne, Australia

La Trobe University, Melbourne, Australia

alan.farley@vu.edu.au

a.farley@latrobe.edu.au

ORCID: 0000-0002-1835-3029

Helen Hong Yang*

La Trobe University, Melbourne, Australia

helen.yang@latrobe.edu.au

ORCID: 0000-0003-1130-6825

**Correspondence*

**Acknowledgement**: This is a part of a larger project funded by Australian Government Office of Learning and Teaching Grant (SD15-5138).

## Abstract

This is the first study that examines the comparative validity of the Chinese National College Entrance Exam (known as Gaokao) to a Western end of secondary school performance measure, Australian Tertiary Admission Rank (ATAR), as admission criteria to an Australian undergraduate degree program. The study offers a strong warning about using just aggregate Gaokao score as a sole admission criterion. However, the findings suggest that with adjustments for gender and time Gaokao is found to be as good an admission criterion as ATAR hence justifying its use for selection purposes if such adjustments are made. Supplementing selection criteria by also using the Gaokao English result is found to be a further improvement. The findings can help inform the admission policy relevant to Chinese international students into undergraduate degree programs globally.



**Keywords**: Admission criteria; Australian Tertiary Admission Rank; Chinese Gaokao; comparative study; validity

## 1. Introduction

The final year Chinese secondary school and university entrance exam, known as the National College Entrance Examination (NCEE or the Gaokao), has become increasingly relevant to Western universities over the last ten years (AEI, 2009; Davey, Lian, & Higgins, 2007; *The Economist*, 2018; Zhang, 2015). This is in line with the continual growth of Chinese international students in Western universities (China's Ministry of Education, 2019). The four major destinations of Chinese school leavers studying overseas, US, UK, Australia, and Canada have different approaches to the use of the Gaokao score as an admission criterion into undergraduate programs. In UK, Gaokao is considered by many UK universities to be at a lower level than those qualifications traditionally used to demonstrate readiness for undergraduate study, with students completing 12 years of education (the Universities and Colleges Admissions Service, UCAS, 2018). In US, there has been a low but growing number of universities accepting Gaokao score (*the Economist*, 2018). The limited acceptance of Gaokao in UK and US universities contrasts with the wide acceptance of Gaokao by universities based in Australia and Canada. Almost all[1] (92%) Australian universities accept Gaokao as a direct admission criterion, and around 30 Canadian universities accept Gaokao score to be used instead of the SAT or similar test (the *Economist*, 2018). The Australian and Canadian universities' admission practice is mirrored by universities in countries such as France, Italy and Spain (Zhang, 2015). Given the different approaches to using Gaokao as an admission criterion to university undergraduate programs across the English speaking countries, the key research question is which admission approach is justified by evidence?

Although inconsistent methodologies are used by Western universities to arrive at decisions about which Gaokao results are suitable for admission, one common underlying assumption implicit in what all these universities do is that Gaokao performance represents a suitable measure of the academic ability needed to succeed in Western universities. This is similar to the way the Australian Tertiary Admission Rank (ATAR) is used to select most school leavers into universities in Australia (see Blyth, 2014). The academic abilities assessed in year 12 programs offered in Australian states, and as summarised by ATAR, are recognised as similar to those assessed in final secondary school years and university admission tests in other countries such as A levels in the UK (PLC, 2015) and the International Baccalaureate Diploma (see VTAC, 2017). Given the recognised similarity between ATAR and some other

[1] The information is collected manually by the authors from Australian universities' websites in 2018.

final year of secondary school/university admission measures of academic ability then a comparison of how the academic abilities assessed by ATAR and Gaokao, in terms of their ability to predict success at university degree level, is a useful indicator of how Gaokao compares to many other final year of secondary school/university admission tests around the world.

Despite the wide acceptance of Gaokao results as an admission criterion in Australian universities, there has not been any study on the comparison between Gaokao and ATAR when used as admission criteria for Australian universities, where the relative validity as admission criteria is assessed by the ability to explain performance in the degree admitted to. This article rectifies that by comparing the explanatory power of Gaokao to that of ATAR, when used as an admission criterion, in predicting the performance of students in an Australian university's undergraduate business degree. The accounting and finance programs in the business degree are selected for the study because they are the most popular business program for Chinese international students in Australian universities (Australian Department of Education and Training, 2018).

Several factors limit the comparative study of Gaokao and any Western countries' tertiary entry tests, as tertiary admission criteria in Western universities. One of the biggest challenges is that the data required for analysis is not easily accessible to researchers. Chinese universities are often not open to allowing access to their student data for research. Adding to the difficulty is the knowledge of tertiary admission systems in the countries compared and the language barrier. The research team of this study comprises members that are capable of reading and writing Chinese at professional level and have a sound understanding of university admission processes in both countries. Researchers were granted access to a unique set of student data of an Australian university's undergraduate business program. Data collection was extensive and complex and took the research team a whole year to complete. One of the strengths of this study lies in its engagement with a comprehensive review of empirical studies on the predictive power of Gaokao for the academic performance in Chinese universities published in Chinese between 1986 and 2017. All Chinese research papers are sourced from China National Knowledge Infrastructure database. The inclusion of Chinese literature aligns with a transformative research paradigm in that the review introduces 'voices' from a country whose research findings on Gaokao (in the Chinese language) have remained unknown in mainstream Western literature. The engagement with Chinese literature will also empower researchers to act in a more informed manner (Yang, Craig, and Farley, 2015). This study offers a useful reference point for future researchers who are interested in comparing Gaokao and the

tertiary admission tests based in other countries that have significant number of Chinese students.

This study offers the first findings on the relativity between Gaokao and a Western university admission criterion (ATAR) as valid admission criteria. Findings of this study will be useful to inform other Western universities' admission policy of students from China where those universities judge their admission criteria for domestic school leavers to be similar to ATAR.

The remainder of the paper is structured as follows. Section 2 provides an overview of the ATAR and the Gaokao. Section 3 presents the research design, and Section 4 discusses the results, limitations and directions for future research.

## 2. ATAR and GAOKAO

### *2.1 Australian Tertiary Admission Rank (ATAR)*

The main selection criterion used for Australian domestic school leavers being admitted to most undergraduate Australian university degrees is a measure called ATAR (students from the State of Queensland[2] are not moving to the ATAR system until 2020, QTAC, 2018). The ATAR[3] of a student is a percentile rank which indicates what percentage of all students of the year 12 age group would be expected to get a score lower than that obtained by the student receiving the ATAR. For example, an individual student with an ATAR score of 90 means the 90% of all other people in the year 12 age group would be expected to perform worse than that student if they completed the year 12 program. This puts the student at the top 10% of the year 12 age group in ATAR performance. Each state in Australia runs a year 12 based program (in many cases as a combined year 11 and 12 program) that produces an aggregate score for each student and then a mathematical process is applied that converts that aggregate score into a ranking. Within each state there is a body that utilises the year 12 results to calculate a measure designed specifically for university selection, for example, in Victoria, the body is called the Victorian Tertiary Admission Centre (VTAC). The year 12 program typically has English as the only compulsory subject and offers a wide range of other subjects. Assessment in these subjects is a mixture of school based tasks and centrally controlled examinations. Each student in each state receives marks for each subject completed along with an aggregate score for their year 12. Each state follows its own procedure for converting students' subject scores (often

[2] In Queensland, the measure is currently called the Overall Position.

[3] See http://www.vtac.edu.au/results-offers/atar-explained.html for further explanation about ATAR

with standardisation and weighting of subject results) into a measure used for university selection. However, they have a number of things in common, one of which is that the end result of each process is a measure that by 2020 will be called ATAR in all states (see VTAC, 2017; QTAC, 2018).

ATAR has long been used as the dominant tertiary admission criterion in Australian universities. It is regarded as a relatively efficient, transparent, and fair tertiary selection method which provides a common measure of students' overall academic ability (Blyth, 2014; Pilcher & Torii, 2018). Prior research based in Australia has reported consistent findings about ATAR being (moderately) positively related to the tertiary academic performance of students in undergraduate subjects (Anderton, 2017; Auyeung & Sands, 1994; Birch & Miller, 2007; Farley & Ramsay, 1988; Messinis & Sheehan, 2015; Wright, 2015). Prior research also reveals that the explanatory power of ATAR differs between individual courses (programs). However, researchers (Baik, Naylor, Arkoudis & Dabrowski, 2017; Craft, 2018; McKenzie & Schweitzer, 2001) caution that ATAR would be more useful when used in conjunction with other measures (e.g. interviews, gender, social and economic status) for admission consideration, rather than ATAR alone.

*2.2 Chinese National College Entrance Examination (the Gaokao)*

In China students undertake an equivalent of the Australian year 12 program called Gaozhong and at the end of this program students who wish to progress to tertiary study sit a set of examinations that form what is called the Gaokao. Bai et al. (2014) provide a useful summary about the Gaokao. The structure of the Gaokao is typically described as "3+X." The three refers to the three compulsory subjects: Chinese, mathematics, and a foreign language (English for the vast majority) for all candidates, with each generally carrying a weighting of 150 marks. The "X" consists of a group of subject tests that differ for students depending on whether they pursue Humanities (H) or Sciences (S) streams. The Humanities stream subjects are history, politics, and geography, and the Sciences stream subjects are physics, chemistry, and biology, with each generally carrying a weighting of 100 marks.

All assessment is exam-based and all the exam papers of the '3+X' component subjects are prescribed by China's Ministry of Education. All exams are closed-book and the exam dates are fixed nationally on June 7 and 8.

The overall Gaokao score for each student is based mainly upon subject scores. However, for a small number of students, the Gaokao score can also include adjustments for factors such

as sports, academic awards, or being children of veterans (Bao & Liu, 2015). Most provinces have a maximum aggregate score of 750.

Compared with the year 12 programs in Australia, the Gaokao program is more rigid due to the limited number of streams and the fixed set of subjects within each stream. The end result of the Gaokao process is individual subject scores of the "3+X" components, and an aggregate score for each student. There is no standardisation of subjects and subject weightings are predetermined.

Prior studies on the predictive power of Gaokao for the academic performance in Chinese universities have generated mixed findings. The courses used in Chinese studies were dominated by science and engineering disciplines. Exceptions are Bai et al. (2014) and Yang (2014) which included business courses. Most Chinese studies found a statistically significant relationship between Gaokao and the academic performance in university courses, except for Bie & Li (1997) and Li (2012). Gaokao score was found to be more related to the first year performance than later years of study in university. The predictive power of Gaokao (aggregate score and Gaokao individual subject results) also varied across the specialisation courses (Bai et al., 2014; Du, Ding, Lin, & Tang, 2016; Gao, 1986; Ning, Xiao, Miao, Dai, & Song, 2001; Wang, Tan, Wang, 2013; Xia, 2017; Yang, 2014). Gaokao English was found to have a stronger predictive power than any other Gaokao individual subject for predicting the academic performance in university (Du et al., 2016; Yang, 2014; Wang et al., 2013; Xie & Zheng, 2009), even though these were courses taught in Chinese. Chinese studies also found gender and admission year had impact on the students' performance (Bai et al 2014; Ning et al., 2001).

## 3. Research Design

### *3.1 The current study*

This study is based on an Australian university's (AU) undergraduate business degree program. The students are admitted to the business degree in Australia based upon ATAR (Australian domestic students, and onshore international secondary students) and Gaokao scores (only a very small number of Chinese international students direct from China). AU also runs an offshore pathway program in conjunction with a Chinese university (CU). The offshore program includes an intensive academic English language program for the first year at CU and AU's first-year business degree program (hereafter FYC) for the second year at CU. A large cohort of Chinese international students is admitted to AU's business degree through this offshore program based upon aggregate Gaokao score. Successful students (i.e. those who pass the FYC, and the embedded English language program) are admitted to the later year of the

degree (in Australia or China) with full credit for first year. For the study period, an average of 80% of the offshore pathway program students transferred to Australia to complete their business degree. The duration of the offshore business degree program in China is four years as opposed to three years in Australia largely due to the intensive English language program in the offshore program. Table 1 illustrates the relationship between AU's onshore and offshore business degree program.

Table 1: AU's onshore and offshore business (accounting and finance) degree program

| AU's onshore program | AU's offshore program in China (CU) | |
|---|---|---|
| | CU Year 1 | English Language Program, concurrent with CU's core subjects |
| AU's first year of degree (FYA) | CU Year 2 | English Language Program, concurrent with AU's first year of degree (FYC) |
| AU's second year of degree | CU Year 3 | AU's second year of degree (to be completed either in China or in Australia) |
| AU's third year of degree | CU Year 4 | AU's third year of degree (to be completed either in China or in Australia) |

*3.2 Research questions*

The review of relevant literature on ATAR and Gaokao reveals similar findings about them being a useful measurement of students' academic ability, despite their limitations. Both Australian and Chinese researchers cautioned the use of ATAR or Gaokao alone as the sole selection criterion of tertiary entry due to the consistently reported moderate or weak predictive power. However, no prior study has investigated the relative validity of Gaokao and ATAR, when used as tertiary selection criteria, on the academic performance of Chinese and Australian students. Hence the specific form of the key research question is:

Research question 1 (RQ1): How do ATAR and Gaokao, if used as the sole tertiary admission criterion, compare in predicting the performance in the first year of the same program?

There are many arguments in the literature and in practice that university admission should not be based upon a single criterion. In practice other criteria are often combined with some aggregate measure of academic ability. In China one such criterion that has consistently been found to be statistically significant in predicting differences in university performance for the same Gaokao score is gender. Another criterion that has been found to be significant in both China and other countries for having incremental predictive power above a basic measure of academic ability is performance in specific subjects/components of an admission test. One such

subject/component of an admission test in China is Gaokao English, which has shown to be significant in predicting university performance even in China when study is in Chinese. Given this study is about admission to an undergraduate degree program taught in English it is justified to assess whether Gaokao English also has incremental predictive power for Western program. Given the study covers multiple admission years the other question that arises is whether any relationship is stable over time. Hence the following supplemental research questions are also examined:

Research question 2 (RQ2): Would adding gender and year variations to ATAR and Gaokao as tertiary admission criteria improve the ability of the admission criteria to predict first year degree performance?

Research question 3 (RQ3): Does adding Gaokao English to the model, as an additional admission criterion, further improve its predictive power?

*3.3 Data collection and study period*

The current study is part of a larger national project funded by the Australian government. A condition of the national grant was that it had prior approval for data collection from both AU and CU. All student data were supplied to the researchers in de-identified form. Three sets of data were available for analysis.

The first set of data was collected for 452 Chinese international students who studied the AU offshore pathway program between 2008 and 2014. For each student data was available on gender, year of admission, aggregate Gaokao, results in each subject in the AU's offshore first year degree (i.e. FYC) and results for all subjects completed in the second and third year degree program, if admitted.

The second set of data was collected for 397 Chinese international students admitted to the first year of AU's offshore program in 2014. It covered aggregate Gaokao score, individual Gaokao English score, gender, and results in each subject in AU's FYC (completed in 2016,the latest data available to the researchers at the time of the study). This is the only dataset with the Gaokao English results (which were supplied by AU from a one-off data collection) and was included solely to allow the testing of RQ3.

The third set of data was collected for 1593 Australian domestic school leavers who studied in the same degree program between 2009 and 2016. For each student data was available on gender, year of admission, ATAR score, and results for all subjects completed in the degree program.

The difference in the start date of 2008 (first data set) versus 2009 (third data set) is explained by the program in China being four years versus three years in Australia (as shown in Table 1).

The descriptive statistics for this data are shown in the Appendix.

### *3.4 Empirical analysis*

#### *3.41 The equivalence of offshore first year and onshore first year degree programs*

Since testing of all three research questions relies upon the equivalence of AU's offshore first year degree in China (i.e. FYC) and the onshore first year of the degree (i.e. FYA). This is the first analysis undertaken. This is done by comparing how well each predicts performance in all second and third year (i.e. later year) subjects in AU's business degree (recognising that 80% of the FYC students studied second and third year degree subjects in Australia).

Two regressions were run using the average mark in all second and third year subjects as the dependent variable in each case (i.e. later year results). The average mark in AU's offshore first year degree (i.e. *FYC*) and average mark in the onshore first year degree (i.e. *FYA*) were used as the two independent variables. Dummy (*D*) variables for year were also included to test if the relationship was stable over time. The 2008 admission year was used as the base year for *FYC* and the 2009 admission year was used for *FYA*. The interpretation of the dummy variables is that they measure the difference in the predicted average later year mark in the designated year versus the base year at each level of FYC or FYA.

The following regression equations (*Eqn.*) were used to test the equivalence of FYC and FYA, and results are provided in Table 2, where LYC and LYA represent the average later year result of students who did the first year of the degree in China (LYC) and Australia (LYA) respectively:

$$Eqn.1: \quad LYC = \alpha + \beta_1 FYC + \beta_2 D_{2009} + \beta_3 D_{2010} + \varepsilon$$

$$Eqn.2: \quad LYA = \alpha + \beta_1 FYA + \beta_2 D_{2010} + \beta_3 D_{2011} + \beta_4 D_{2012} + \beta_5 D_{2013} + \varepsilon$$

Table 2: Regression results of *Eqn. 1* and *Eqn. 2*

| | Unstandardised Coeff. Eqn. 1 LYC | Standardised Coeff. Eqn. 1 LYC | Unstandardised Coeff. Eqn. 2 LYA | Standardised Coeff. Eqn. 2 LYA |
|---|---|---|---|---|
| Constant | 9.766** | | -0.877 | |
| Average mark in offshore first year degree (FYC) | 0.755** | 0.748** | | |
| Average mark in onshore first year degree (FYA) | | | 0.910** | 0.741** |
| 2009 Dummy | -0.531 | -0.034 | | |
| 2010 Dummy | -0.692 | -0.044 | -0.806 | -0.044 |
| 2011 Dummy | | | -4.294 | -1.465 |
| 2012 Dummy | | | -0.828 | -0.768 |
| 2013 Dummy | | | -0.870 | -0.665 |
| Adjusted R squared | 0.548 | 0.548 | 0.551 | 0.551 |
| F statistic | 178.444** | 178.444** | 51.784** | 51.784** |
| No. of Observations | 440 | 440 | 209 | 209 |

* represents coefficients significant at the 5% level
** represents coefficients significant at the 1% level

The models have very similar high explanatory power of the variation in later year results (54.8% for FYC and 55.1% for FYA) with similar standardised coefficients for the impact of each one standard deviation change in average first year result on average later year result (0.748 versus 0.741 standard deviations change in later year result). Both models show no significant instability in the relationship over time. The results in Table 2 offer sufficient evidence of the equivalence of the academic knowledge level achieved in offshore FYC compared to that achieved in onshore FYA in predicting later year performance (See Section 4.4 for further discussion). This allows the study to treat them as equivalent in the following analysis of the three research questions.

*3.42 RQ1: How do ATAR and Gaokao, if used as the sole tertiary admission criterion, compare in predicting the performance in the first year of the same program?*

To address RQ1 a series of regressions were run. The dependent variable was either the average result in AU's first year degree subjects taught onshore (i.e. *FYA*) or the average result in AU's offshore first year degree subjects taught in China (i.e. *FYC*). The results of these models are shown in Table 3 and are based upon the following regression equations:

$$Eqn.3: \quad FYA = \alpha + \beta_1 ATAR + \varepsilon$$

$$Eqn.4: \quad FYC = \alpha + \beta_1 Gaokao + \varepsilon$$

Table 3: Regression results of *Eqn 3* and *Eqn 4*

| | Unstandardised Coefficients Eqn. 3 FYA | Standardised Coefficients Eqn. 3 FYA | Unstandardised Coefficients Eqn. 4 FYC | Standardised Coefficients Eqn. 4 FYC |
|---|---|---|---|---|
| Constant | 41.119** | | 44.644** | |
| ATAR | 0.326** | 0.459** | | |
| Gaokao | | | 0.068** | 0.154** |
| Adjusted R squared | 0.209 | 0.209 | 0.021 | 0.021 |
| F statistic | 196.940** | 196.940** | 10.863** | 10.863** |
| No. of Observations | 741 | 741 | 452 | 452 |

* represents coefficients significant at the 5% level
** represents coefficients significant at the 1% level

The coefficient results in Table 3 show that a one percentile point change in ATAR score is expected to lead to a 0.326 change in average first year result (or a one standard deviation change in ATAR data is expected to lead to a 0.459 standard deviation change in average first year mark) whereas a one aggregate mark change in aggregate Gaokao (out of 750) is expected to increase average first year mark by 0.068 (or a one standard deviation change in aggregate Gaokao score is expected to increase average first year mark by 0.154 standard deviations).

While all models (Eqn 3 and Eqn 4) in Table 3 are statistically significant, they do not support Gaokao alone being comparable to ATAR alone as an admission criterion. This is due to the fact that Gaokao alone only explains 2.1% of the variation in the average first year results. In comparison, ATAR alone explains 20.9% of the variation in average first year results. The power of ATAR to explain performance at university is known to drop as the ATAR score drops. Since AU's intake has the majority of its domestic students with ATARs from the fifth and sixth deciles, 20.9% is in line with expectations from other studies (e.g. Messinis & Sheehan, 2015).

*3.43 RQ2: Would adding gender and year variations to ATAR and Gaokao as tertiary admission criterion improve the ability of the admission criteria to predict first year degree performance?*

To answer RQ2, variables representing gender and year were added (using 2009 as the base year for ATAR and 2008 as the base year for Gaokao) to improve the explanatory power

of admission criteria. This is because both variables have been found to be significant in studies of Gaokao's predictive power of performance in Chinese universities (see Section 2.2), and gender could be easily incorporated into admission criteria while year allows for variation in standards between years. The results are reported in Table 4 and are based upon the following regression equations:

$$Eqn.5:\ FYA = \alpha + \beta_1 ATAR + \beta_2 FEMALE + \beta_3 D_{2010} + \beta_4 D_{2011} + \beta_5 D_{2012} + \beta_6 D_{2013} + \varepsilon$$

$$Eqn.6:\ FYC = \alpha + \beta_1 Gaokao + \beta_2 FEMALE + \beta_3 D_{2009} + \beta_4 D_{2010} + \varepsilon$$

Table 4: Regression results of *Eqn. 5* and *Eqn. 6*

| | Unstandardised Coefficients Eqn. 5 FYA | Standardised Coefficients Eqn. 5 FYA | Unstandardised Coefficients Eqn. 6 FYC | Standardised Coefficients Eqn. 6 FYC |
|---|---|---|---|---|
| Constant | 44.079** | | 30.731** | |
| ATAR | 0.295** | 0.415** | | |
| Gaokao | | | 0.083** | 0.189** |
| Female | 0.585 | 0.034 | 6.544** | 0.400** |
| 2009 Dummy | | | 1.391 | 0.086 |
| 2010 Dummy | -0.926 | -0.047 | 3.687** | 0.227** |
| 2011 Dummy | -1.570 | -0.037 | | |
| 2012 Dummy | -2.319** | -0.100** | | |
| 2013 Dummy | -4.619** | -0.190** | | |
| Adjusted R squared | 0.237 | 0.237 | 0.218 | 0.218 |
| F statistic | 39.263** | 39.263** | 32.400** | 32.400** |
| No. of Observations | 741 | 741 | 452 | 452 |

* represents coefficients significant at the 5% level
** represents coefficients significant at the 1% level

The results from Table 4 suggest that for the same Gaokao score female students on average would expect to achieve an average result in AU's offshore first year degree program 6.544 marks higher than the male students. This lends support to the view that in China female students perform better than male students in tertiary undergraduate degree programs for the same Gaokao score, or viewed another way, are discriminated against or disadvantaged in the Gaokao process. The significance of both gender and Gaokao when used together suggests that Gaokao correctly ranks the academic ability of each of the males and females separately but a correct ranking in a pooled group is only achieved when the Gaokao scores of females are adjusted upwards. The standardised coefficients show that gender has greater impact than

Gaokao, highlighting the importance of the gender adjustment. However, the statistical insignificance of the Female variable in the equation 5 model suggests gender adjustment is not supported for Australian domestic students.

Table 4 also indicates that the relationship is not stable over time for either of the ATAR or Gaokao models since at least one year dummy is significant in each model. This can be explained by either the standard of the ATAR or Gaokao changing over time or the standard set in the degree changing over time, or some combination of both. With respect to ATAR it is unlikely that the scale of ATAR changes over such a short period since as a percentile rank measure that can only occur if the distribution of academic ability within the year 12 school age population changes over such a short time frame. With Gaokao being an aggregate score measure it is more possible that its standard can change from year to year because it depends directly on the standard of the examinations set. To eliminate this risk aggregate Gaokao would need to be converted to a percentile rank distribution measure in a manner similar to ATAR. The standard set in the degree can certainly alter from year to year since it is directly related to the standard set on individual assessment tasks.

*3.44 RQ3*: *Does adding Gaokao English to the model, as an additional admission criterion, further improve its predictive power?*

To answer RQ3, the second dataset was used for analysis. The result in the individual Gaokao English subject was added to the model since selection criteria could also be modified to include this. For this cohort student data was available on gender, Gaokao (aggregate) score (*GA*), results in FYC, and Gaokao English score (*GE*). The results are shown in Table 5, and are based upon the following regression equation:

$$Eqn.7: \quad FYC = \alpha + \beta_1 GA + \beta_2 GE + \beta_3 FEMALE + \varepsilon$$

Table 5: Regression results of *Eqn. 7*

| | Unstandardised Coefficients Eqn. 7 FYC | Standardised Coefficients Eqn. 7 FYC |
|---|---|---|
| Constant | 44.310** | |
| Gaokao Aggregate (GA) | -0.005 | -0.028 |
| Gaokao English (GE) | 0.204** | 0.246** |
| Female | 7.700** | 0.412** |
| Adjusted R squared | 0.273 | 0.273 |
| F statistic | 50.678** | 50.678** |
| No. of Observations | 395 | 395 |

* represents coefficients significant at the 5% level
** represents coefficients significant at the 1% level

Results from Table 5 suggest Gaokao English had greater explanatory power than Gaokao aggregate score. As indicated in Table 5, adding Gaokao English increased the percentage of variation in first year results explained from 21.8% to 27.3%. However, what was unexpected was that aggregate Gaokao had no incremental predictive power beyond Gaokao English as shown by the Gaokao coefficient being statistically insignificant.

## 4. Discussion

### *4.1 Comparison of the predictive power of ATAR and Gaokao, when used as the sole admission criterion, for academic performance*

The findings of this study (Table 3) reveal that the predictive power of ATAR and Gaokao are not comparable when used as the sole admission criterion for academic performance. One possible explanation of the difference in results can be due to the type of measure ATAR is versus Gaokao. Gaokao is a raw aggregate score that generally follows something like a negatively skewed normal distribution (Ashour & Abdel-hameed, 2010) whereas ATAR is a percentile rank distribution that converts a similarly distributed raw aggregate score into a uniform distribution. As a robustness test of whether this was the cause of this result the Gaokao scores were converted into a percentile rank distribution equivalent and then the Gaokao regression was rerun. The explanatory power remained almost unchanged. Hence, this was not the cause of the difference.

### *4.2 Comparison of the predictive power of ATAR and Gaokao, when used along with gender, as admission criteria, for the academic performance*

The results from Table 4 suggest that one reason for the poor explanatory power of Gaokao alone is that equal Gaokao scores for males and females do not represent equal levels of the

forms of academic ability needed to succeed in Western universities. The finding about the adding gender as an adjustment to Gaokao is consistent with the prior research based in Chinese university programs (e.g. Bai et al., 2014), and some studies based in Western countries (e.g. Voyer & Voyer, 2014). From the perspective of Western universities' admission policy, the finding of this study indicates that if they make an upward adjustment to the Gaokao scores of females when considering them for admission this would yield similar validity as a predictor of performance as using ATAR as an admission criterion. However, as indicated in Table 4, in this study there is no support for using gender as an admission criterion for domestic students at AU, consistent with many other studies based in Western universities (e.g. Guney, 2009).

It is acknowledged that politically it may be difficult to justify using gender to differentiate one admission pathway and not others but this approach is strongly supported by almost all research on Gaokao along with much of the research on ATAR. Once both gender and year effects are taken into account the predictive power of the set of admission criteria including Gaokao improves to be almost as good as the set including ATAR, and close to a level found for ATAR as a predictor of accounting students' performance in other studies based in Australia (e.g. Farley & Ramsay, 1988). This puts the explanatory power of Gaokao, with suitable adjustments, almost on a par with the major measure (ATAR) used to select domestic students in Australian universities.

*4.3 Adding Gaokao English as an additional admission criterion*

The finding that Gaokao English had a greater explanatory power than Gaokao aggregate score is consistent with the studies based in Chinese universities (e.g. Du et al., 2016). One explanation for the Chinese findings is the need for reading English language literature is particularly relevant to the fields of Chinese Engineering and Commerce degrees, resulting from the internationalisation of Chinese higher education. This study found the aggregate Gaokao score had no incremental predictive power beyond Gaokao English. This finding indicates that not only did Gaokao English contain useful information about English language proficiency, but also was at least as good a measure of relevant academic abilities as aggregate Gaokao. Hence, findings of this study suggest that selection could be improved even further if instead of using aggregate Gaokao, Western universities use Gaokao English as part of the selection criteria, or likely a more acceptable option, using aggregate Gaokao with bonus points based on Gaokao English, after making a similar adjustment for gender. The results in this study suggest the above approach would have greater validity as a set of admission criteria for Chinese students than using ATAR, as currently done with domestic students.

This study offers a strong warning about using just aggregate Gaokao score as an admission criterion and hence, it supports the UK and US university reticence to admit students on this basis. However, with appropriate selectiveness and adjustment, the study shows that Gaokao data can be used as a suitable measure for admission directly into degrees in Western universities and hence if used appropriately would support the approach of Australia and Canada (and some other countries) provided the key additional factors are considered (gender and preferably Gaokao English).

The Chinese language literature cited in this study highlights that the key findings about Gaokao in this article are not limited to its use for admission to English language Western universities but generally apply equally to its use for admission to Chinese universities.

*4.4 Limitations and directions for future research*

The validity of this study could be improved upon if the moderated equivalence of first year programs (as shown by the results from Table 2) was replaced by actual equivalence of the first year programs by utilising a cohort of students who have been admitted to the onshore first year program based upon their Gaokao results. All students would then attend the same classes taught by the same staff and complete exactly the same assessment, rather than relying on a moderation process to create the equivalence. This is not an option at AU, who have very few such students, but is an extension that could be conducted at some other Australian or Canadian universities.

The validity of the results involving Gaokao English would also be improved if a sample were available where both the first year results for students admitted via Gaokao, with Gaokao English results, and students admitted via ATAR were available in the same year.

Relevance of the study could also be extended by covering a wider range of degree programs, including those outside of business, with Engineering being another major area of study for Chinese students. However, at AU there are very few Chinese undergraduate students pursuing disciplines other than Accounting and Finance.

A further issue of validity of the study is the extent to which universities in Western countries other than Australia require similar academic skills and knowledge to those in Australia. This study has shown that Gaokao is a measure of the skills and knowledge required to succeed in Australian universities. This only translates to other Western countries if their universities have similar requirements. The fact that some universities in other Western countries accept ATAR as a measure of the skills and knowledge required for admission suggests there is some alignment between what is required in Australia and what is required in

other Western countries. There is also the more general question of what such measures are being used for at university admission. Are they about measuring the level of skills and knowledge or are they more about assessing the general academic ability of potential students? To the extent they are used as a general measure of academic ability at higher levels of study the alignment of skill and knowledge development between final secondary programs/university entrance tests becomes less relevant as long as they produce a similar ranking with respect to academic ability.

The development of a learner's academic ability is an evolving process, and thus changing over the life of the tertiary study. Therefore, it would be appropriate for university policy makers to consider a principled approach to admission decisions which involves weighing the attributes of applicants and the academic requirements of the applicant's future discipline specific study.

**References**


Anderton, R., 2017, Identifying factors that contribute to academic success in first year allied health and science degrees at an Australian University, *Australian Journal of Education*, 61(2), 184–199

Ashour, S.K. & Abdel-hameed, M.A., 2010, Approximate skew normal distribution, *Journal of Advanced Research*, 1, 341–350

Australian Education International, 2009, Research on China's National College Entrance Examination (the Gaokao), https://www.austrade.gov.au, accessed 1 May 2015.

Auyeung, P. K., & Sands, D., 1994, Predicting success in first-year university accounting using gender-based learning analysis. *Accounting Education*, 3(3), 259-272.

Bai C.E, Chi W., & Qian X.Y., 2014, Do college entrance examination scores predict undergraduate GPAs? A tale of two universities, *China Economic Review*, 30, 632-647

Baik, C., Naylor, R., Arkoudis, S., & Dabrowski, A., 2017, Examining the experiences of first-year students with low tertiary admission scores in Australian universities, *Studies in Higher Education*, 1-13.

Bao, R. & Liu, N.N., 2015, Analysis of Gaokao bonus point policies since China's Reform, *Journal of Higher Education Management*, 9(5), 30-35. (In Chinese)

Bei, XJ. & Li, Z.C., 1997, Statistical analysis of the relationship between Gaokao results and university performance, *Jiancai Higher Education Theory and Practice*, 4, 49-52. (In Chinese)

Birch, E. R., & Miller, P. W., 2007, Tertiary entrance scores: Can we do better? *Education Research and Perspectives*, 34(2), 1-23.

Blyth, K., 2014, Selection methods for undergraduate admissions in Australia. Does the Australian predominate entry scheme the Australian Tertiary Admissions Rank (ATAR) have a future?, *Journal of Higher Education Policy and Management*, 36(3), 268-278.

Craft, J. A., 2018, Academic performance of low SES students at an Australian university satellite campus, *Studies in Higher Education*, 1-14.

Davey, G, De Lian, C & Higgins, L., 2007, The university entrance system in China, *Journal of Further and Higher Education*, 31(4), 385-396.

Du, XY., Ding, HC., Lin, XF., & Tang, H., 2016, Study on the correlation between the Gaokao score and the first year degree performance, *Journal of Anhui University of Technology(Social Sciences)*, 33(4), 56-57. (In Chinese)

*Economist* (United Kingdom), 2018, Some Western universities see merit in China's flawed exam. *The Economist*, 414 (9095).

Farley, A. & Ramsay, A, 1988, Student performance in first year tertiary accounting courses and its relationship to secondary accounting education. *Accounting & Finance* 28(1), 29-44.

Gao, S.Z., 1986, Correlation analysis of Gaokao and university performance, *Chongjing Normal College Academic Journal (Natural Sciences)*, 1(3), 53-57. (In Chinese)

Guney, Y., 2009, Exogenous and endogenous factors influencing students' performance in undergraduate accounting modules. *Accounting Education, 18*(1), 51-73.

Li, H., 2012, Correlation analysis of medical students' Gaokao results and their university performance, *Qiqihar Medical School Academic Journal*, 33(14), 1922-1923. (In Chinese)

McKenzie, K., & Schweitzer, R., 2001, Who Succeeds at University? Factors predicting academic performance in first year Australian university students, *Higher Education Research & Development,* 20(1), 21-33.

Messinis, G., & Sheehan, P., 2015, The academic performance of first year students at Victoria University by entry score and SES, 2009–2013., https://www.ncsehe.edu.au, accessed 20 April 2018

Ministry of Education, China, 2019, Statistics of Chinese students going overseas study in 2018, http://www.moe.gov.cn, accessed 11 May 2019

Ning, J., Z., Xiao, J., Miao, B.Q., Dai, X.L., & Song, C.N., 2001, *Higher Education Science Education*, 37(3), 46-50. (In Chinese)

Pilcher, S. & Torii, K., 2018, Crunching the number: Exploring the use and usefulness of the Australian Tertiary Admission Rank (ATAR), Mitchell Institute paper No. 01/2018, www.mitchellinstitute.org.au, accessed 10 January, 2019

Presbyterian Ladies' College Sydney, 2015, Top British Universities, https://www.plc.nsw.edu.au/microsites/global-education/studying-overseas/studying-in-the-uk, accessed 20 May, 2019

QTAC, 2018, Queensland Tertiary Admission Center Queensland ATAR 2020, https://www.qtac.edu.au/atar-my-path/atar, accessed 11 July 2018

UCAS, 2018, China Gaokao, https://qips.ucas.com/qip/china-gaokao, accessed 10 May 2019

Voyer, D., & Voyer, S. D., 2014, Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140(4), 1174-1204.

VTAC, 2017, ABC of Scaling, http://www.vtac.edu.au/files/pdf/publications/abcofscaling.pdf, accessed 13/04/2018

Wang, C.J., Tan, C.C., & Wang., H, 2013, Statistical analysis of the Gaokao results and the academic results of university students, *College Mathematics*, 29(4), 79-86. (In Chinese)

Wright, V. J., 2015, Is ATAR useful for predicting the success of Australian students in initial teacher education? *Australian Journal of Teacher Education*, 40 (9), 1–15.

Xia, D.M., 2017, Correlation analysis of humanity students' Gaokao results and university performance, *Journal of Anqing Normal University (Natural Science Edition)*, 23(2), 53-62. (In Chinese)

Xie. Z.C. & Zheng, H. J., 2009, Correlation analysis of Gaokao and the academic results in university, *Mathematics in Practice and Theory*, 39(12), 1-6. (In Chinese)

Yang, G.L., 2014, Are all admission sub-tests created equal? — Evidence from a National Key University in China, *China Economic Review*, 30, 600–617.

Yang, H.H., Craig, R., & Farley, A., 2015, A review of Chinese and English language studies on corporate environmental reporting in China. *Critical Perspectives on Accounting*, 28, 30-48.

Zhang, HZ., 2015, Italy, France accept Gaokao scores, *China Daily Europe*, June 26, http://europe.chinadaily.com.cn/china/2015-06/26/content_21110695.htm, accessed on 10 April 2018

Appendix

Descriptive statistics

| Measure | Dataset 1 | Dataset 2 | Dataset 3 |
|---|---|---|---|
| Percent of Females | 67 | 75 | 51 |
| Mean Gaokao Score | 498 | 537 | |
| Min Gaokao Score | 435 | 345 | |
| Max Gaokao Score | 568 | 593 | |
| Mean Gaokao English Score | | 122 | |
| Min Gaokao English Score | | 90 | |
| Max Gaokao English Score | | 143 | |
| Mean ATAR Score | | | 59.0 |
| Min ATAR Score | | | 30.2 |
| Max ATAR Score | | | 99.3 |